Statistics Tutorial (Part 1 of 2)

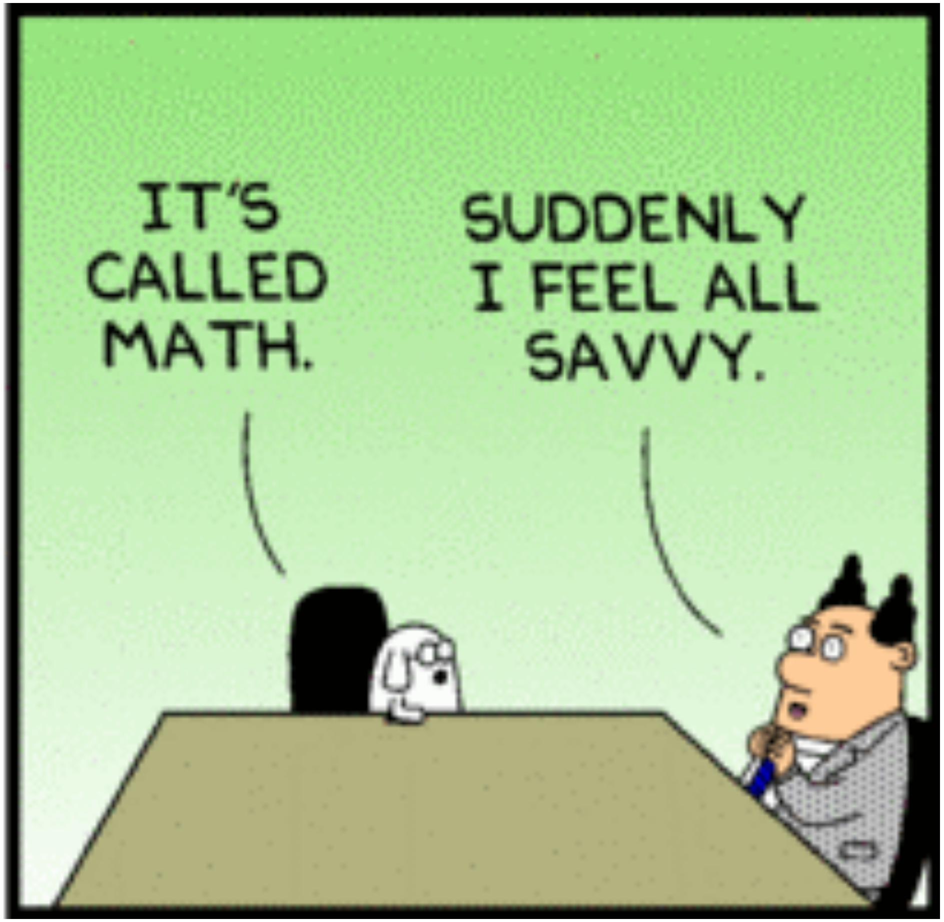
The mark of a truly educated man is to be moved deeply by statistics.

-- George Bernard Shaw (1856 – 1950)
Irish playwright and winner of the Nobel Prize for Literature (1925)

Dogbert's risk management advice

Dogbert's risk management advice

Dogbert's risk management advice



(Just about) all the statistics you'll need

- In this lecture. . .
  - discrete and continuous probability distributions
  - expected value, variance, standard deviation, covariance, and correlation
  - Numerical examples – expected returns and risks for 2-asset portfolios

Discrete & Continuous Distributions

- There are two types of random variables: discrete and continuous.
  - **Discrete random variables** can only take on a finite number of "countable" values; e.g., time and temperature rounded to the nearest minute or degree.
  - **Continuous random variables** can take on an infinite number of possible values; e.g., the non-rounded versions of time and temperature.

Discrete Probability Distribution

A discrete probability function is a function that satisfies the following properties:

1. The probability that the random variable X can take a specific state contingent value X_s is $p(X_s)$; that is, $\Pr[X = X_s] = p(X_s) = p_s$.
2. p_s is non-negative for all possible values of X_s.
3. The sum of p_s over all possible states is 1; i.e.,

$$\sum_{s=1}^{n} p_s = 1.$$

A consequence of properties 2 and 3 is that $0 \leq p_s \leq 1$ for all s.

Continuous Probability Distribution

The mathematical definition of a continuous probability function, $f(x)$, is a function that satisfies the following properties.

1. The probability that x is between two points a and b is $\Pr[a \leq x \leq b] = \int_a^b f(x)dx$.
2. $f(x)$ is non-negative for all possible values of x.
3. The integral of the probability function is one; that is, $\int_{-\infty}^{\infty} f(x)dx = 1$.

Expected value

- Expected value is also known as the mean, or average value for a random variable; it represents the *central value* about which variable observations scatter.
- If you roll a die, there is an equal (1/6) probability of each number coming up (note: since all outcomes are equally probable, this is an example of a "uniform" probability distribution).
- What is the expected value or average number if you roll the die many times? The answer is $1 \times (1/6) + 2\times(1/6) + 3\times(1/6) + 4\times(1/6) + 5\times(1/6) + 6\times(1/6) = 3\frac{1}{2}$.

Expected value

- Generally, if we have a random variable X (the number thrown, say)
- which can take any of the values X_s (in this case, 1,2,3,4,5, or 6) for $s = 1, \ldots, N$
- each of which has a probability p_s (in this case, 1/6)
- then the expected value is $E(X) = \sum_{s=1}^{n} p_s X_s$.

Properties of expected values

- Expected values have the following properties:
  - $E(c) = c$. (The expected value of a constant is the constant).
  - $E(cX) = cE(X)$ (The expected value of a constant times a random variable is equal to the constant multiplied by the expected value of the random variable).
  - $E[X + Y] = E[X] + E[Y]$ (The expected value of a sum of random variables is equal to the sum of the expected values of the random variables).

Variance and Standard Deviation

- Variance is the expected value of the squared deviation of the random variable from its mean.
- Standard Deviation is the square root of the variance and it measures how far most of the variable observations scatter about the mean.
- Standard Deviation is commonly used in finance and risk management as a definition for risk (although as a risk measure, it has a number of important shortcomings that we'll discuss at a future time!).

Variance and Covariance

- The variance $Var(X)$ is computed as follows:

$$Var(X) = \sigma_X^2 = E[(X - E(X))^2] = \sum_{s=1}^{n} p_s (X_s - E(X))^2.$$

  - The standard deviation measures the spread, or dispersion around the expected value.
  - In the case of the toss of a die, the standard deviation is 1.71 (check this for yourself!).
- Variances have the following properties:
  - $Var(cX) = c^2 Var(X)$ (The variance of a constant times a random variable is equal to a constant squared times variance).

Variance and Covariance

- Variance properties (continued):
  - If X and Y are statistically *independent*, $Var(X + Y) = Var(X) + Var(Y)$. (Total variance is the sum of variances)
  - If X and Y are statistically *dependent*, $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$. (Total variance is sum of variances & covariances)
- Covariance between X and Y is computed as follows:

$$Cov(X,Y) = \sigma_{xy} = E[(X - E(X))(Y - E(Y))]$$

$$= \sum_{s=1}^{n} p_s (X_s - E(X))(Y_s - E(Y)).$$

- Correlation coefficient: $\rho_{XY} = Cov(X,Y)/\sigma_X \sigma_Y$.
  - Correlation is a "standardized" covariance
  - Covariance is defined over the open interval (-∞, +∞), whereas correlation is defined over the closed interval [-1, +1].

Statistics Class Problem

Suppose the return distributions for two risky assets are as follows:

State	p_s	$r_{a,s}$	$r_{b,s}$
1	1/3	-3%	36%
2	1/3	9%	-12%
3	1/3	21%	12%

1. Calculate the expected returns for assets *a* and *b*.
2. Calculate the variances and standard deviations for assets *a* and *b*.
3. Calculate the covariance and correlation between assets *a* and *b*.
4. Calculate the expected return and standard deviation for an equally weighted portfolio consisting of asset *a* and *b*.
5. Determine the least risky combination of assets *a* and *b* and calculate the expected return and standard deviation for such a portfolio.

Minimum Risk Portfolio (2 assets)

- Previously, we calculated the expected return and standard deviation for *an equally weighted* portfolio consisting of securities *a* and *b*.
- Next, let's determine a security weighting scheme which minimizes portfolio risk. In other words, what combination of w_a and w_b accomplishes this objective?

Minimum Risk Portfolio (2 assets)

$$\sigma_p^2 = w_a^2\sigma_a^2 + w_b^2\sigma_b^2 + 2w_a w_b\sigma_{ab}$$

$$= w_a^2\sigma_a^2 + (1-w_a)^2\sigma_b^2 + 2w_a(1-w_a)\sigma_{ab}$$

$$= w_a^2\sigma_a^2 + (1-w_a)(1-w_a)\sigma_b^2 + 2w_a\sigma_{ab} - 2w_a^2\sigma_{ab}$$

$$= w_a^2\sigma_a^2 + \sigma_b^2 + w_a^2\sigma_b^2 - 2w_a\sigma_b^2 + 2w_a\sigma_{ab} - 2w_a^2\sigma_{ab}$$

$$= w_a^2(\sigma_a^2 + \sigma_b^2) + 2w_a(\sigma_{ab} - \sigma_b^2) + \sigma_b^2 - 2w_a^2\sigma_{ab}.$$

Therefore,

$$\frac{d\sigma_p^2}{dw_1} = 2w_a(\sigma_a^2 + \sigma_b^2) + 2(\sigma_{ab} - \sigma_b^2) - 4w_a\sigma_{ab}$$

$$= w_a(\sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}) + \sigma_{ab} - \sigma_b^2 = 0$$

$$\Rightarrow w_a = \frac{\sigma_b^2 - \sigma_{ab}}{\sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}}.$$

Minimum Risk Portfolio (2 assets)

- See http://fin4335.garven.com/fall2019/Two Asset Portfolio.xls